Bureaucratic Quality and Electoral Accountability

TARA SLOUGH   *New York University, United States*

*I*n many theories of electoral accountability, voters learn about an incumbent's quality by observing public goods outcomes. But empirical findings are mixed, suggesting that increasing the visibility of these outcomes only sometimes improves accountability. I reconcile these heterogeneous findings by highlighting bureaucrats' role in the production of public goods. In a simple model of electoral accountability involving a voter, a politician, and a bureaucrat, I show that accountability relationships yield distinct empirical implications at different levels of bureaucratic quality. To illustrate how this model rationalizes otherwise mixed or heterogeneous results, I develop a new research design—a theoretically structured meta-study—to synthesize existing findings. Meta-study evidence on the accountability of Brazilian mayors suggests that a common model of electoral accountability that allows for variation in bureaucratic quality predicts observed heterogeneity in politician and voter behavior and beliefs across multiple studies with distinct samples, treatments, and outcomes.

Under what conditions are incumbent politicians induced to provide public goods for their constituents? When can voters harness their collective voice at the ballot box to select or incentivize politicians to deliver public goods? The vast literature on electoral accountability that responds to these questions is characterized by widespread heterogeneity and mixed findings (e.g., Enríquez et al. 2023; Incerti 2020). I argue that consideration of bureaucrats' role in the production of public goods provides a new explanation for observed variation in the manifestations of politician–voter accountability relationships globally.

Politicians worldwide rely on bureaucrats to co-produce public goods and services. Politicians typically appropriate funds to a public goods project—like a road—while bureaucrats ensure that the road is built. Yet, politicians in different contexts must rely on bureaucracies of markedly different qualities. Whereas a politician facing a high-quality, capable bureaucracy may be highly confident that bureaucrats will ensure that the road is competently constructed, a politician facing a lower quality bureaucracy is likely less confident that the project will be capably executed. Bureaucratic quality, therefore, conditions the efficiency of politicians' investments in public goods, which, in turn, shapes politicians' incentives to invest in these public goods over private rents. Variation in politicians' investment strategies affects what voters can learn from observing public goods and, thus, the extent to which voters can use information to retain or select high-performing politicians or motivate politicians to invest in public goods. Distinct manifestations of electoral accountability, therefore, emerge at different levels of bureaucratic quality. The observable implications of these equilibria can reconcile mixed or heterogeneous empirical findings about electoral accountability.

The theory that I advance contrasts with recent explanations for observed variation in voter–politician accountability relationships across the world's democracies. Many recent contributions emphasize uninformed or irrational voters as a barrier to functioning electoral accountability relationships. In such accounts, citizens who lack access to an informational signal about a politician's type do not learn about the politician's quality (i.e., Dunning et al. 2019; Ferraz and Finan 2008; Humphreys and Weinstein 2012). Alternatively, citizens may receive information but fail to rationally update (Achen and Bartels 2016; Healy, Malhotra, and Mo 2010).[1] In either case, failure to access or update on performance-relevant information breaks the link between a politician's actions and her subsequent electoral fortunes, so the politician maximizes her own utility, regardless of whether their actions are congruent with voters' preferences. I instead show that these patterns can emerge with informed, rational (Bayesian) voters due to bureaucratic co-production of public goods.

To understand the implications of my argument about the importance of bureaucratic quality, I develop a simple two-period model of electoral accountability with a bureaucrat. As is standard, voters evaluate politicians on the basis of observed policy outcomes, here levels of public goods provision (Ashworth 2012). However, the model departs from standard practice by considering co-production of public goods by a politician and an unelected bureaucrat. The politician allocates a budget to public goods or private rents (corruption). In turn, bureaucrats produce public goods with the allocated funds. Bureaucratic quality influences the efficiency with which these public goods are

Tara Slough [ID], Assistant Professor, Department of Politics, New York University, United States, tara.slough@nyu.edu.

---

[1] But see Fowler and Montagnes (2015), Fowler and Hall (2018), and Ashworth, Bueno de Mesquita, and Fridenberg (2018).

produced. Politicians differ in their competence at getting things done, which translates into their ability to monitor bureaucrats. Voters update their beliefs about a politician's competence on the basis of public goods outputs. The voter then decides whether or not to retain the incumbent for a second period or to elect a challenger.

When bureaucratic quality is low, it is inefficient for either type of politician to fund public goods. As a result, both competent and incompetent politicians pool on corrupt behavior, allocating all funds to private rents instead of public goods. Due to this pooling, voters are not able to ascertain politician quality, regardless of whether they observe the lack of public goods (the signal). At moderate levels of bureaucratic capacity, public goods provision is efficient for competent types but inefficient for incompetent types. In this case, voters' observation of public goods allows for updating on politician type, which yields higher retention of competent types. At higher levels of bureaucratic quality, partially pooling and then pooling equilibria emerge in which both types invest in public goods with positive probability. As in the separating equilibrium, in these equilibria, voter observation of public goods outputs permits updating, thereby facilitating more frequent retention of competent politicians.

These theoretical results have implications for how we interpret the results of empirical studies of electoral accountability. Suppose that researchers conduct an experiment in which they provide voters with information on an incumbent's corruption (the politician's behavior). At low levels of bureaucratic quality, information that the incumbent is corrupt should not have an effect on voters' beliefs about the politician or their voting behavior because both types of politicians pool by allocating the budget to rents. The pooling equilibria at high levels of bureaucratic quality further suggests that voters' beliefs should not change in response to revelation of clean (non-corrupt) politician behavior since both types of politicians pool on allocating the budget to public goods.[2]

I illustrate the plausibility of the theory by proposing a *theoretically structured meta-study* research design to synthesize existing work on electoral accountability. I examine studies of accountability of Brazilian mayors. Specifically, I develop and validate a measure of bureaucratic quality across Brazilian municipal bureaucracies. I use this measure to extend important studies of corruption by politicians (Avis, Ferraz, and Finan 2018; Ferraz and Finan 2011); voter updating on politician performance (Weitz-Shapiro and Winters 2016; Winters and Weitz-Shapiro 2016); and the emergence of incumbency disadvantage for local mayors (Klasnja and Titunik 2017). I find that equilibria in which competent politicians always invest in public goods and incompetent politicians may invest in public goods in a first term, only to shirk by appropriating rents in a

second term, emerge only in municipalities with low levels of bureaucratic quality relative to the sample of Brazilian municipalities. In municipalities with higher quality bureaucracies, politicians abstain from appropriating private rents, politicians do not shirk in their second term, and voters do not update on a signal of a clean politician. These findings are consistent with the implications of a pooling equilibrium that emerges at high levels of bureaucratic quality.

This article makes theoretical, substantive, and methodological contributions. The theory connects accountability models focused on a voter and politician to models of moral hazard in bureaucracies. Joining Yazaki (2018), Li, Sasso, and Turner (2023), and Foarta (2023), the model posits distinct roles for politicians and bureaucrats in the production of public goods as central to our understanding of electoral accountability. The integration of bureaucratic quality in a model of electoral accountability generates new insights about the relationship between outcomes during an incumbent's tenure and their re-election prospects. For example, Ashworth, Bueno de Mesquita, and Fridenberg (2018) show how outcomes outside an incumbent's control may affect the incumbent's re-election prospects. I show, conversely, that even in a such world with Bayesian voters, outcomes under a politician's control may have limited or no effect on the incumbent's re-election prospects at some levels of bureaucratic quality because of pooling behavior by politicians of different types. Further, the theory engages a recent argument by Martin and Raffler (2021) that joint policymaking by politicians and bureaucrats hinders voter learning about politicians. By endogenizing the behavior of politicians and bureaucrats, the present results reveal that their finding that bureaucratic co-production limits voter learning is not general across all levels of bureaucratic quality. Instead, when bureaucratic quality is high, reliance on bureaucrats to produce public goods *improves* voters' information, thereby promoting the positive selection of politicians.

Substantively, the results engage large empirical literature on information and accountability in developing democracies (e.g., Bhandari, Larreguy, and Marshall 2021; Chong et al. 2015; Cruz, Keefer, and Labonne 2021; Dunning et al. 2019). The theory rationalizes mixed findings with respect to the (average) effects of information on voter beliefs and election outcomes (Enríquez et al. 2023). Moreover, the empirical results show different instantiations of electoral accountability in the Brazilian context, including incumbent rent-seeking (Ferraz and Finan 2008; 2011), voter learning (Boas and Hidalgo 2011; Weitz-Shapiro and Winters 2016; Winters and Weitz-Shapiro 2016), and incumbency disadvantage (Klasnja and Titunik 2017), vary systematically in bureaucratic quality and can be explained by a common model.

Beyond the study of bureaucracy and electoral accountability, a central contribution of this article is to use applied theory to integrate results in established literatures with heterogeneous or mixed findings, thereby providing a novel method for the accumulation of evidence. Traditional meta-analysis rests on an

---

[2] In contrast, if the researchers provided information about public goods outputs rather than politician behavior, the voter could update on the politician's type at high levels of bureaucratic quality.

assumption that constituent studies estimate the same quantity or empirical target (Slough and Tyson 2023). But these approaches are limited in how they address heterogeneity in effects[3] and are not coherently applicable in contexts where constituent studies have conceptually distinct treatments or outcomes. By introducing theoretically structured meta-studies, I show how researchers can accumulate evidence in contexts or literatures with distinct treatments and outcomes, but in which these varied treatments and outcomes can be reasonably represented as constituent pieces of a common theoretical model.

THEORY

Consider three actors: an incumbent politician, P; a bureaucrat, B; and a voter, V. In each of two periods, the politician and bureaucrat jointly produce public goods. After first-term public goods production, there is an election in which the politician contests re-election against a challenger.

Politicians are of an incompetent or competent type, $\theta \in \{\underline{\theta}, \overline{\theta}\}$, respectively. The politician's type is private information to the politician and the bureaucrat. The voter holds a prior belief that the politician is a competent type with probability $Pr(\theta = \overline{\theta}) = \pi \in (0, 1)$. I conceive of competence as ability to manage the bureaucracy or get things done via oversight. A competent politician monitors the bureaucrat at intensity $\overline{\theta}$, while an incompetent politician monitors the bureaucrat at intensity $\underline{\theta}$, where $0 < \underline{\theta} < \overline{\theta} < 1$. This formulation of politician competence for overseeing the bureaucracy follows closely from recent formulations of oversight capability or capacity (Martin and Raffler 2021; Raffler 2022), while emphasizing that these capabilities vary across individual politicians as in Brierly (2020).

Public goods are produced as a function of the funding allocated by the politician in period t and the quality and effort of the bureaucracy. Specifically, politicians allocate a budget, normalized to 1 in each period, to public goods (a_t) or private rents ($1 - a_t$). The quality of the bureaucracy, $q > 1$, is exogenous and common knowledge. While bureaucratic quality may be an outcome of policies pursued by a politician, the model simply assumes that quality is slow-moving and requires sustained investment to realize changes (Huber and Ting 2021; Rauch 1995).[4] Recent evidence by Besley et al. (2022) suggests that, globally, bureaucratic quality has been remarkably persistent within countries since at least 1900. Treating bureaucratic quality as common

knowledge reflects the fact that the modal citizen–government interaction is between citizens and bureaucrats, providing citizens ample opportunity to observe bureaucratic quality. Further, there are generally many more bureaucrats than politicians, meaning that many citizens are likely to know bureaucrats in both official and unofficial capacities.

I assume that the bureaucrat exerts costly effort, e in response to some intensity of oversight, given by $\theta \in \{\underline{\theta}, \overline{\theta}\}$. As such, the utility of the bureaucrat, in period t, net of a wage satisfying their participation constraint, can be written:

$$u_t^B(e) = -\theta(1 - e_t) - \frac{e_t^2}{2}. \tag{1}$$

Note that θ is given by the politician's type. The bureaucrat is not forward-looking and chooses a level of effort in each period. This characterization of a bureaucrat is intentionally very simplistic as the focus of this article is to draw out implications for the voter–politician accountability relationships.

Given the allocation of funds by the politician and the effort exerted by a bureaucrat, the public good, $g_t(a_t, e_t)$, is produced according to the production function:

$$g_t(a_t, e_t) = \begin{cases} a_t q, & \text{with probability } e_t, \\ 0, & \text{with probability } 1 - e_t. \end{cases} \tag{2}$$

This production function indicates that if the politician invests in public goods ($a_t > 0$), then the expected quantity of public goods outputs is increasing in bureaucratic quality (q) and effort (e_t). In contrast, if the politician starves public goods funding ($a_t = 0$), they are not produced. The production function in Equation 2 further clarifies the relationship between bureaucratic quality, q, and broader notions of bureaucratic capacity. Bureaucratic capacity consists of both the skill of bureaucrats (Geddes 1994), their allocation of bureaucrats across a jurisdiction (Acemoglu, García-Jimeno, and Robinson 2015), and the effort exerted by bureaucrats. I capture the first two features in quality (q) and the third in bureaucratic effort (e_t). To the extent that bureaucratic capacity is one key input to broader notions of state capacity, bureaucratic quality should be viewed as one constituent component of state capacity.[5] In the present framework, both bureaucratic quality and effort increase the efficiency with which a politician's funding allocation is converted to a public goods output.

The politician trades off private rents for public goods when allocating the budget. Both types of politicians value the provision of public goods. The assumption that both types to value public goods provision is

---

[3] Meta-regression methods do not require a common empirical target, but specify a known structural relationship between empirical targets that accounts for heterogeneity.

[4] There is limited empirical evidence on reforms designed to improve bureaucratic quality. Notably, however, interventions studied in existing literature that intend to improve bureaucratic quality via hiring (selection) were initiated by higher levels of government from *outside* the localities they serve, not by local politicians who may be judged on the quality of their services (Ashraf et al. 2020; Dal Bó, Finan, and Rossi 2013).

[5] Conceptualizing of a state as an organization consisting government (politicians and bureaucracy) and civil society clarifies that state capacity is broader than bureaucratic capacity (Berwick and Christia 2018).

not common to all accountability models. Note, however, that many of the qualitative results of the model can be generated by a venal/non-venal politician type space (instead of competence).[6] I note these departures when discussing results. Competent types' superior ability to induce the bureaucrat to work is captured in the realization of g_t in the per-period utility of the (incumbent) politician in Equation 3:

$$u_t^P(a_t; \theta) = 1 - a_t + g_t. \quad (3)$$

The politician receives $u_t^P(a_t; \theta)$ for each period she is in office, and utility normalized to 0 if she is not in office. This normalization creates a re-election incentive for the politician. As such, the politician's utility over two periods is given by

$$u^P(a_1, a_2; \theta) = \begin{cases} 2 - a_1 - a_2 + g_1 + g_2, & \text{if re-elected,} \\ 1 - a_1 + g_1, & \text{if not re-elected.} \end{cases} \quad (4)$$

The voter may observe the realization of first-term public goods provision, and in turn forms a posterior belief about the politician's type, μ. To understand the role of voter information—here, whether or not a voter observes the public goods—in generating results, I assume that the voter observes a signal, $z = g_1$, of first-term public goods with probability $p \in [0, 1]$ and that the voter does not receive a signal, $z = \varnothing$, with complementary probability $(1-p)$. The idea that some voters might passively observe public goods outputs is straightforward. Voters routinely observe the state of roads or public transportation when they travel about a community. They may observe the state of local schools or clinics when they send their children to school or seek public health services. These quotidian observations do not require deep familiarity with government procedures or utilization of right to information laws. For this reason, I treat public goods outputs as the informational signal in the baseline model.

While it is natural to think that incumbents, challengers, or civil society would publicize performance signals, the information and accountability literature generally assumes some barrier to diffusion of this information—for instance, a lack of local media (Ferraz and Finan 2008; Larreguy, Marshall, and Snyder 2020). The assumption of exogenous revelation solely maintains that these diffusion technologies are not manipulated by politicians in the short run. Further, treating p as exogenous is consistent with the theoretical treatment of information revelation in experiments on information and accountability (Izzo, Dewan, and Wolton 2022). In this article, setting $p = 0$ allows for characterization of equilibria in which voters are (completely) uninformed (see anecdotes throughout Dunning et al. 2019).

The voter values consumption of the public good. This implies that the voter cares about a politician's competence whenever competent politicians produce more public goods (in expectation). The voter's utility is thus given by expected public goods provision in the second period and a valence shock for the incumbent, parameterized as $\phi \sim U[-b, b]$, where $b > q$. The voter votes, $v \in \{i, c\}$, to re-elect the incumbent (i) or elect the challenger (c). If elected, a challenger acts as a first-period incumbent. This assumption facilitates the study of the effect of electoral incentives on politician behavior in settings with a two-term limit.[7] For that reason, I index second-period actions by i and c, respectively. The voter's second-period expected utility from the a vote for incumbent or a vote for a challenger (c), can be expressed:

$$E[u_2^V(i)] = E[g_2^i | z] + \phi, \quad (5)$$

$$E[u_2^V(c)] = E[g_2^c]. \quad (6)$$

In evaluating $E[g_2^i | z]$ and $E[g_2^c]$, the voter considers differences in expected politician competence and differences in allocation behavior that depend on whether a politician is term limited.

Sequence and Equilibrium Concept

The game proceeds according to the sequence:

1. Nature determines θ, the incumbent's competence. Only the incumbent and bureaucrat observe θ.
2. The incumbent allocates a_1 to the public good.
3. The bureaucrat exerts effort e_1 to produce the first-term public good, g_1.
4. With probability p, the voter observes $z = g_1$ and forms a posterior belief about the politician's type, μ. The valence shock ϕ is revealed, and the voter chooses whether to re-elect the incumbent or elect the challenger.
5. If the incumbent was re-elected, she allocates a_2^i to the public good. Otherwise, the challenger allocates a_2^c to the public good.
6. If the incumbent was re-elected, the bureaucrat exerts effort e_2^i to produce the public good g_2^i. Otherwise, the bureaucrat exerts effort e_2^c to produce the public good g_2^c.

I characterize the Perfect Bayesian Equilibria (PBE) of the game. The incumbent's allocation decision is the choice $a_1 \in \{0, 1\}$. The bureaucrat's effort allocation is $e_1 \in \mathbb{R}_+$. Public goods production, $g_1 : \{0, 1\} \times \mathbb{R}_+ \to \{0, q\}$, maps the budget allocation and bureaucratic

---

[6] Differing objectives by venal and non-venal politicians could also capture differences in politicians' incentives to monitor the bureaucracy as emphasized by Raffler (2022) since public goods require co-production and private rents do not.

[7] Alternatively, the assumption that challengers act as first-period incumbents approximates an infinite-horizon model with term limits in which voters solely look to the next term when casting their votes.

effort into a public goods output observed by all players. Voters update beliefs on the observation of public goods $\mu : \{0, q\} \rightarrow [0, 1]$ and the voter's voting strategy is a mapping $v : \{0, q\} \times [0, 1] \rightarrow \{i, c\}$. The second-period incumbent's allocation strategy is a mapping $a_2^i : \{0, q\} \times [0, 1] \times \{i, c\} \rightarrow \{0, 1\}$. Finally, second-period bureaucratic effort and public goods production represents the mapping: $e_2^i : \{0, q\} \times [0, 1] \times \{i, c\} \times \{0, 1\} \rightarrow \mathbb{R}_+$ and public goods provision represents the mapping $g_2^i : \{0, q\} \times [0, 1] \times \{i, c\} \times \{0, 1\} \times \mathbb{R}_+ \rightarrow \{0, q\}$. As in many signaling games, there exist multiple equilibria in some regions of bureaucratic quality. I invoke the intuitive criterion refinement to ensure that the equilibria discussed below are unique (at each level of bureaucratic quality) (Cho and Kreps 1987).

EQUILIBRIUM ANALYSIS

First, consider the bureaucrat's equilibrium level of effort. By straightforward inspection of the bureaucrat's objective, it is clear that optimal effort, $e_t^* = \theta$. The bureaucrat's effort depends only on the politician's type in either period. When combined with Equation 2, this optimal effort indicates that politician competence and bureaucratic quality are complements with respect to the production of public goods.[8] This contrasts with the idea that politician type (quality) and bureaucratic quality are substitutes, which is motivated by the observation that high-quality bureaucracies tend to insulate outputs from the follies of bad politicians. Instead, the model develops an alternate mechanism for this observed insulation focused on how bureaucratic quality shapes a politician's allocation decision.

Turning to the incumbent's second-term allocation strategy, the politician considers the expectation of second-term public goods provision, $E[g_2^i(a_2^i, e_2^i)] = \theta q a_2^i$. Where $E[g_2^i(a_2^i, e_2^i)] \geq 1$, a politician will invest the budget in public goods, $a_2^i = 1$. In contrast, where $E[g_2^i(a_2^i, e_2^i)] < 1$, a politician will invest nothing, $a_2^i = 0$. The politician's optimal second-period allocation strategy is given by

$$a_2^{i*} = \begin{cases} 1, & \text{if } q \geq \frac{1}{\theta}, \\ 0, & \text{else .} \end{cases} \tag{7}$$

Intuitively, if the bureaucracy is of sufficiently low quality (low q), neither type has an incentive to fund public goods because it is inefficient to do so. This implies that even a competent politician that values public goods outputs will take private rents when the bureaucracy is incapable of efficiently producing public goods. On the other hand, when q is sufficiently high, both types will fund public goods. The efficiency gains in the provision of public goods from a high-quality

---

[8] This complementarity can be seen from the fact that $\frac{\partial^2 E[g_t(a_t, \theta_t)]}{\partial q \partial \theta_t} = a_t \geq 0$.

bureaucracy induce both types of politician to fund public goods, insulating outputs (to some extent) from incompetent politicians.

Consider the voter's voting decision. The voter votes for the incumbent if $E[u_2^V(i)] > E[u_1^V(c)]$. Given the distribution of the valence shock and denoting equilibrium allocation strategies for each term and politician type as $\boldsymbol{a}$, the incumbent's probability of victory is

$$\begin{aligned} \tau(\mu, \boldsymbol{a}) &= \frac{1}{2} + \frac{E[g_2^i | z] - E[g_2^c]}{2b} \\ &= \frac{1}{2} + \frac{\mu E[g_2^i(a_2^i, e_2^i | \theta = \overline{\theta})] + (1-\mu) E[g_2^i(a_2^i, e_2^i | \theta = \underline{\theta})]}{2b} - \\ &\quad \frac{\pi E[g_2^c(a_2^c, e_2^c | \theta = \overline{\theta})] + (1-\pi) E[g_2^c(a_2^c, e_2^c | \theta = \underline{\theta})]}{2b}. \end{aligned}$$
(8)

Turning to the voter's beliefs and voting decision, recall that the voter observes $z = g_1$ with probability p. With probability $1-p$, the voter does not observe public goods outputs ($z = \varnothing$). Consider the latter case first. In this case, $\mu = \pi$, which follows (trivially) from Bayes' rule. If voters do not update, a politician's re-election fate is independent of her first-period allocation decision. If this occurs, the politician maximizes her utility by adopting the same allocation strategy in both periods, always adopting the optimal allocation strategy given by Equation 7. Thus, following Equation 8, the probability of re-election is $\tau(\pi, \boldsymbol{a}) = \frac{1}{2}$.

When voters do observe first period public goods, $z \in \{0, q\}$, they are able to update their beliefs on the basis of observed outputs. However, at different levels of bureaucratic quality, the signal offered by the realization of public goods differs in its informativeness. Politicians choose their first period allocation behavior on the basis of efficiency considerations and their anticipated prospects for re-election. Define four cutpoints in q that are relevant for characterizing the resultant equilibria: $q_1 \equiv \frac{1}{\overline{\theta}}$, $q_2 \equiv \max\{\frac{1}{\underline{\theta}}, \frac{2b(1-\pi\overline{\theta})}{\underline{\theta}(2b(1-\pi\overline{\theta}) + \overline{\theta}p(1-\pi))}\}$, $q_3 \equiv \frac{2b(\theta(\pi-1)-\overline{\theta}\pi)(1+\underline{\theta}(\pi-1)-\overline{\theta}\pi)}{\underline{\theta}\left(2b(\theta(\pi-1)-\overline{\theta}\pi)(1+\underline{\theta}(\pi-1)-\overline{\theta}\pi) + \overline{\theta}(\overline{\theta}-\underline{\theta})p(\pi-1)\pi\right)}$, and $q_4 \equiv \frac{1}{\underline{\theta}}$. It is straightforward to see that $q_1 \leq q_2 \leq q_3 \leq q_4$ and that $q_1 < q_4$ under the parametric assumptions of the model.
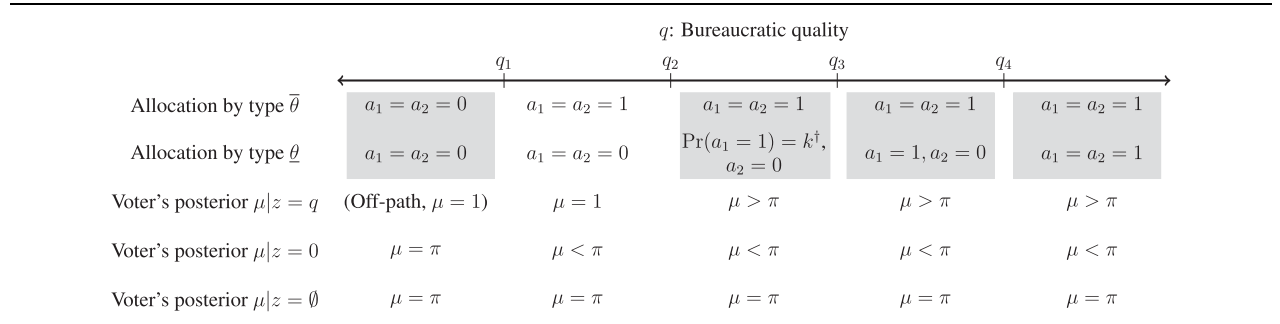
Proposition 1 (Equilibrium)

In the unique PBE:
(i) If $q < q_1$, both types of politicians allocate $a_1 = a_2 = 0$ to public goods.
(ii) If $q \in [q_1, q_2)$, a competent-type politician allocates $a_1 = a_2 = 1$, while an incompetent-type politician allocates $a_1 = a_2 = 0$ to public goods.
(iii) If $q \in [q_2, q_3)$, a competent-type politician allocates $a_1 = a_2 = 1$, while an incompetent-type politician allocates $a_1 = 1$ with probability $k \in (0, 1)$, $a_1 = 0$ with probability $1-k$, and $a_2 = 0$ to public goods.
(iv) If $q \in [q_3, q_4)$, a competent-type politician allocates $a_1 = a_2 = 1$, while an incompetent-type politician allocates $a_1 = 1$ and $a_2 = 0$ to public goods.

FIGURE 1.   Equilibrium Allocation Behavior by Politicians and the Voter's Posterior Belief as a Function of Bureaucratic Quality, q

	q: Bureaucratic quality				
		q_1	q_2	q_3	q_4
Allocation by type $\bar{\theta}$	$a_1 = a_2 = 0$	$a_1 = a_2 = 1$	$a_1 = a_2 = 1$	$a_1 = a_2 = 1$	$a_1 = a_2 = 1$
Allocation by type $\underline{\theta}$	$a_1 = a_2 = 0$	$a_1 = a_2 = 0$	$\Pr(a_1 = 1) = k^\dagger, a_2 = 0$	$a_1 = 1, a_2 = 0$	$a_1 = a_2 = 1$
Voter's posterior $\mu\|z = q$	(Off-path, $\mu = 1$)	$\mu = 1$	$\mu > \pi$	$\mu > \pi$	$\mu > \pi$
Voter's posterior $\mu\|z = 0$	$\mu = \pi$	$\mu < \pi$	$\mu < \pi$	$\mu < \pi$	$\mu < \pi$
Voter's posterior $\mu\|z = \emptyset$	$\mu = \pi$	$\mu = \pi$	$\mu = \pi$	$\mu = \pi$	$\mu = \pi$

Note: The shaded rectangles represent pooling or partially pooling equilibria in which both types of politicians make the same allocation in the first period with positive probability. †In the partially pooling equilibrium, the incompetent type allocates $a_1 = 1$ with probability k and $a_1 = 0$ with probability $1-k$. The Supplementary Material provides an explicit derivation of k.

(v) If $q \geq q_4$, both types of politicians allocate $a_1 = a_2 = 1$ to public goods.

Figure 1 depicts both politician allocations and voter beliefs in each equilibrium region. In multiple regions of the parameter space, competent and incompetent politicians pool (or partially pool) by making the same first-period allocation. Importantly, this allocation behavior by politicians determines how much voters can learn by observing public goods.

When bureaucratic quality is low ($q < q_1$), neither type of politician invests in public goods. If the voter observes the signal that no public goods were produced ($z = 0$), they do not learn anything because both types of politicians divert the budget to rents. This produces the same outcomes—in terms of voter beliefs and behavior—that we would expect if voters never saw the signal ($z = \emptyset$). Indeed, in this pooling equilibrium, voters do not care about the politician's type because neither type will allocate funds to or produce the public good.

In the remaining equilibria, the voter can learn if they observe the public goods signal, as is apparent in Figure 1. As is standard, when the voter sees that public goods were produced ($z = q$), they update positively such that $\mu > \pi$. When the voter sees no public goods produced ($z = 0$), they update negatively such that $\mu < \pi$. Because re-election rates are increasing in the voter's posterior belief that the incumbent is competent, this learning means the voter re-elects competent types at a strictly higher rate than incompetent types (see Lemma A1 in the Supplementary Material). This generates positive selection of second-period incumbents. Voters stand to learn more in the separating equilibrium than the partially pooling or pooling equilibria that emerge at higher levels of bureaucratic quality.[9] In this separating equilibrium, an observation

that public goods were produced ($z = q$) reveals the incumbent to be a competent type. Importantly, however, even in the two pooling equilibria in which both types of politician allocate first-period funds to public goods, voters can also update on the politician's type in order to retain competent politicians at a higher rate. Here, the complementarity between politician competence and bureaucratic quality allows voters to update on the basis of observed public goods.

Three attributes of the partially pooling and pooling equilibria that emerge in the region $q \in [q_2, q_4)$ are important. First, note that these equilibria only emerge when the voter could be informed (when $p > 0$).[10] If the politician knows the voter cannot observe public goods, the incompetent type has no incentive to pool with the competent type in the first period because it will not improve their re-election chances. Second, we should only observe an *increase* in shirking by the politician in the second period in these equilibria. In all other equilibria, each type of politician makes the same public goods allocations in each period, so term effects can only be driven by the positive selection of second-term incumbents. This term effect on shirking is clearly driven by incompetent types who are re-elected. Third, these equilibria feature an incumbency *disadvantage*. Within the present model, incumbency disadvantage presents when the unconditional probability of re-election of first-term politicians falls below $\frac{1}{2}$, as shown in Corollary A2 in the Supplementary Material.[11] Incumbency disadvantage is driven by voter anticipation of second-term shirking. It is costly for voters to re-elect an incompetent type for a second term when either type of politician would allocate the budget to public goods (with positive probability) in their first term. This anticipation leaves the voter more likely to elect a first-term challenger than re-electing the incumbent for a second term. This explanation for

---

[9] The pooling equilibrium that emerges for $q \geq q_4$ would not emerge with a venal/non-venal type space (as opposed to competence). With such a type space, it is possible to generate equilibria that are qualitatively similar to the three equilibria that emerge at any $q < q_4$.

[10] This is apparent from inspection of q_2, q_3, and q_4. When $p = 0$, $q_2 = q_3 = q_4$.
[11] The unconditional re-election rate is given by $\pi\Pr(\text{re}-\text{elected}|\theta = \bar{\theta}) + (1-\pi)\Pr(\text{re}-\text{elected}|\theta = \underline{\theta})$.

FIGURE 2. Equilibrium Allocation Behavior by Politicians and the Voter's Posterior Belief as a Function of Bureaucratic Quality, q, in Extension with Observable Politician Behavior

	q: Bureaucratic quality				
	$\tilde{q}_1$	$\tilde{q}_2$	$\tilde{q}_3$	$\tilde{q}_4$	
Allocation by type $\overline{\theta}$	$a_1 = a_2 = 0$	$a_1 = a_2 = 1$	$a_1 = a_2 = 1$	$a_1 = a_2 = 1$	$a_1 = a_2 = 1$
Allocation by type $\underline{\theta}$	$a_1 = a_2 = 0$	$a_1 = a_2 = 0$	$\Pr(a_1 = 1) = k^{\dagger},$ $a_2 = 0$	$a_1 = 1, a_2 = 0$	$a_1 = a_2 = 1$
Voter's posterior $\mu \vert z' = 1$	(Off-path, $\mu = 1$)	$\mu = 1$	$\mu > \pi$	$\mu = \pi$	$\mu = \pi$
Voter's posterior $\mu \vert z' = 0$	$\mu = \pi$	$\mu = 0$	$\mu < \pi$	(Off-path, $\mu = 0$)	(Off-path, $\mu = 0$)
Voter's posterior $\mu \vert z' = \emptyset$	$\mu = \pi$	$\mu = \pi$	$\mu = \pi$	$\mu = \pi$	$\mu = \pi$

Note: The shaded rectangles represent pooling equilibria in which both types of politicians make the same allocation in the first period with positive probability. [†]In the partially pooling equilibrium, the incompetent type allocates $a_1 = 1$ with probability k and $a_1 = 0$ with probability $1 - k$. The Supplementary Material provides an explicit derivation of k.

incumbency disadvantage is analogous to the mechanism in Klasnja and Titunik (2017), though the present theory introduces new scope conditions—with respect to bureaucratic quality—on where we should observe it.

Extension: Observable Politician Behavior

Many studies of information and accountability emphasize provision of information about a politician's actions—here, their allocation behavior—rather than information about public goods outputs. This information is frequently sourced from audits by higher levels of government (e.g., Arias et al. 2022; Ferraz and Finan 2008; Larreguy, Marshall, and Snyder 2020), data collected by civil society organizations (e.g., Humphreys and Weinstein 2012), or researcher codings of otherwise obscure public information (e.g., Adida et al. 2020). Relative to public goods signals, voters are less likely to observe politicians' allocation behavior directly on a day-to-day basis, precisely because this information is technical, obscure, or obscured.

Nevertheless, given the frequent use of information about politician behavior in recent research on electoral accountability, I now describe a model extension in which voters observe politician allocation behavior, a_1, instead of public goods outputs, g_1, and politicians make allocations in anticipation of the voter (potentially) observing this behavior. Voter observation of politician allocations instead of public goods outcomes removes a source of randomness: whether or not public goods are successfully produced by the bureaucrat as in Martin and Raffler (2021). I characterize the PBE of this game in Proposition A1 in the Supplementary Material and plot equilibrium allocation behavior and voter beliefs in Figure 2. I denote the thresholds in bureaucratic quality in this extension as $\tilde{q}_1 = \frac{1}{\overline{\theta}}, \tilde{q}_2, = \max\{\frac{1}{\overline{\theta}}, \frac{2b}{\underline{\theta}2b + p\overline{\theta}}\}$ $\tilde{q}_3 = \frac{2b}{\underline{\theta}2b + p\overline{\theta}\pi}$, and $\tilde{q}_4 = \frac{1}{\underline{\theta}}$.

While the five equilibria are similar to those in Proposition 1 and Figure 1, the differences are informative. Consistent with the findings of Martin and Raffler (2021), in the separating equilibrium, co-production with the bureaucrat adds noise to the signal. As a consequence, the voter learns more in a separating equilibrium when they observe politician actions, not public goods. But note that $\tilde{q}_2 \leq q_2$, which means that the separating equilibrium is sustained for a smaller range of the bureaucratic quality parameter space when voters observe politician actions, not outputs. This occurs because the incompetent type begins to (partially) pool with the competent type to increase their likelihood of re-election at lower levels of bureaucratic quality.

The pooling equilibria that emerge for any $q \geq \tilde{q}_3$ provide different implications for voter welfare. In these equilibria, both politicians pool by allocating $a_1 = 1$ to public goods. In these cases, the voter cannot learn anything from observing the politician's action because both types make identical allocations. Indeed, a signal that a politician allocated funds to public goods to generates no updating by the voter. As a result, in these regions, the voter would be better off observing public goods outputs. Here, the complementarity between politician competence and bureaucratic quality ensures that competent politicians deliver the public goods at higher rates than incompetent politicians. As such, in contrast to the findings in Martin and Raffler (2021), for sufficiently high bureaucratic quality, co-production of public goods between politicians and bureaucrats *facilitates* rather than hinders electoral accountability.

This extension suggests that there are important welfare effects of the type of informational signals that voters use to update their beliefs about politicians. If studies of information and accountability overemphasize signals consisting of politician actions relative to the public goods signals that voters routinely observe, the literature may mischaracterize the welfare consequences of voter information for public goods provision, selection of politicians, and politicians' electoral incentives. Moreover, because bureaucratic quality determines which type of information—public goods or politician actions—is more advantageous for voters to access, our assessments of voter welfare may be systematically biased (in different directions) at different levels of bureaucratic quality.

Voters might also try to learn about an incumbent's competence by observing bureaucrats' behavior when they seek services. However, this would require voters to observe bureaucratic effort. Effort is generally assumed to be unobservable (or at least not directly observed). Indeed, the politician's inability to observe bureaucratic effort generates the bureaucrat's moral hazard problem in the present model. A second extension in Appendix D1.1 of the Supplementary Material analyzes this case, which is arguably less central to quotidian accountability relationships than the public goods signal and less central to academic research on accountability than the politician action signal.[12]

ACCOUNTABILITY OF BRAZILIAN MAYORS

I present empirical evidence in support of this theory by developing a new research design, a theoretically structured meta-study. This design seeks to synthesize the findings of multiple studies that measure different empirical implications of a common equilibrium (resp. equilibria). I use the theory to generate predictions about how measures of politician allocation behavior, voter beliefs, and voter behavior vary systematically in bureaucratic quality.

I consider the case of municipal governance in Brazil, which is arguably the most researched case in the study of electoral accountability of local politicians. Since the work of Ferraz and Finan (2008), researchers have studied mayoral corruption and voters' responses to revelations of corruption or lack thereof (Ferraz and Finan 2011; Weitz-Shapiro and Winters 2016; Winters and Weitz-Shapiro 2016). These articles use the results —real or hypothetical—of federal audits of municipalities to understand politician behavior and voter learning. In separate work, Klasnja and Titunik (2017) find evidence of a strong incumbency disadvantage in Brazilian mayoral elections. None of these works consider how bureaucratic quality may condition these findings. I draw upon the above theory to synthesize these sets of findings that employ different treatments and outcomes in different samples of municipalities and elections. To do so, I leverage a new measure of bureaucratic quality and extend these analyses.

Measuring Bureaucratic Quality

I measure bureaucratic quality at the municipal level. As the theory clarifies, bureaucratic quality (q) is distinct from bureaucratic effort (e) and public goods outputs (g). The measure of bureaucratic quality should therefore abstract from effort or public goods outputs. As above, I conceptualize quality as a measure of human capital of individuals employed in municipal administration. I rely on Brazil's Basic Municipal Information Survey (MUNIC) to measure characteristics of

employment in (direct) municipal administration. This survey, implemented by the Instituto Brasiliero de Geografia e Estadística, requires municipalities to report counts of public employees working in direct municipal administration, disaggregated according to several categories including education and contract type. Given that the raw data consist of counts of public employees, the level of cross-sectional aggregation is the municipality.

I operationalize bureaucratic quality as the average education level of bureaucrats working in municipal administration. The measure of bureaucratic quality capture features of a representative (average) public employee. I abstract from measures of the number of public employees per capita for two reasons. First, per-capita measures of municipal employment do not account for efficiencies of scale: running a fixed set of programs requires more employees per capita in small municipalities. Second, one form of clientelism that is likely to affect bureaucratic composition is the use of patronage jobs to reward donors or party supporters. Classic descriptions of patronage in Latin America include accounts of low-wage workers expanding out the ranks of public employment (Calvo and Murillo 2004; Grindle 2012). This is generally believed to degrade bureaucratic quality. Both concerns show why per-capita or count-based measures introduce artifacts aside from the human capital of bureaucrats.

In treating bureaucratic education as a measure of quality, several legal, economic, and political considerations are warranted. Legally, municipal employees in direct administration should be hired with civil service provisions, though empirically adherence varies substantially, with many municipalities relying heavily on contractors. Variation in hiring practices is substantial across municipal governments, which accords with the wide observed variation in bureaucratic quality (e.g., Toral 2023).

Average bureaucratic education clearly is driven, in part, by local labor market conditions. The scope of heterogeneity across Brazilian labor markets is likewise impressive. I account for regional variation using state fixed effects. I also use flexible covariate specifications to adjust for municipal population, average municipal education (years of education), formality (percentage of workers working in the formal sector), and per-capita GDP. All of these features correlate with the measure of bureaucratic quality (Supplementary Figure A2). However, they (collectively) account for less than 20% of the variation in bureaucratic quality, indicating that this variation is not simply a function of variation in local labor markets (Supplementary Figure A3).

Analysis of persistence of this measure of bureaucratic education within municipalities over five waves of MUNIC (2005, 2008, 2011, 2014, and 2018) accords with some qualitative assumptions of the model. While quality is secularly increasing over time (Supplementary Figure A1), within municipalities bureaucratic quality is sticky. The annualized autocorrelation of bureaucratic quality between waves of the survey is 0.84 (Supplementary Table A3). I further show that

---

[12] Since equilibrium bureaucratic effort reveals the politician's type, variation in politician allocation strategies is purely driven by efficiency considerations in this extension.

such persistence obtains across the constituent education categories. In the model, q is treated as exogenous and the politician does not alter bureaucratic quality. The data are consistent with this assumption. Supplementary Table A4 reports the results of first-difference models that regress changes in bureaucratic quality (from consecutive waves of MUNIC) on indicators measuring a change in mayor and change in the mayor's party in an intervening election. I find no consistent evidence that changes in mayor or mayor's party yield differential shifts in average bureaucratic quality: point estimates are near-zero and precisely estimated. Further evidence from visualization of the ECDFs of changes in bureaucratic quality shows no evidence of differential changes in variance (Supplementary Figure A4).

Finally, given the importance of distinguishing the implications of the present theory of bureaucratic quality and accountability from theories premised only on voter information, I examine the association between bureaucratic quality and local media presence in Supplementary Table A5. In the Brazilian municipal context, the presence of community radio is argued to be the most important form of media for diffusing local news (Boas and Hidalgo 2011; Ferraz and Finan 2008; Varjão 2019). I show that while raw measures of bureaucratic quality and radio station presence are positively correlated, conditional on the local labor market covariates and state fixed effects, the conditional association between bureaucratic quality and radio presence is estimated to be a precise zero. I use an indicator for radio presence as an additional covariate to allay possible concerns.

Measuring Politician Allocation Behavior

In this analysis, I measure politicians' allocations to rents, $1-a$, through the results of federal audits of municipal governments. Such audits have gained prominence across Latin America in the last two decades and have been used to measure corruption in academic literature (Chong et al. 2015; Ferraz and Finan 2008). In Brazil, the audits I use to measure allocations are conducted by the federal Controladoria-Geral da União (CGU), through a municipal auditing program inaugurated in 2003. Because municipalities in Brazil receive the majority of their budgets from the federal government, such audits cover sizable shares of municipal budgets.[13] Audits consist of visits by a team of federal officials to municipalities to oversee allocation and disbursement of funds and observe outputs. They report their findings in reports which are disseminated by local media (Ferraz and Finan 2008).

Brazil's audits are oft-studied as a natural experiment because the federal government randomly selects municipalities by lottery. While this random assignment has facilitated many studies of the effects of audits (i.e., Avis, Ferraz, and Finan 2018; Ferraz and Finan

2008), from the perspective of this article, it provides random sampling of municipalities, which ensures support across all levels of bureaucratic quality. Support across the distribution of bureaucratic quality is essential to making inferences about the theory proposed by this article.

A final consideration about the use of audit data to measure allocations considers whether audits measure the actions of politicians, here, mayors, or the municipal administration generally. I follow existing studies of accountability in attributing corrupt or malfeasant spending to mayors. This is precisely the inference that experimental studies of accountability ask voters to make (Arias et al. 2022; Boas, Hidalgo, and Melo 2019; Chong et al. 2015). In Brazilian municipalities, executives are responsible for proposing a budget and, alongside the city council, monitoring its execution (Gonçalves 2013). It is not surprising, therefore, that opposition to audits in the form of lawsuits have come from elected politicians, not other municipal officials (Seabra 2018). Furthermore, studies of the effects of audits find few consequences for bureaucrats, at least in terms of retention (Ferrali and Kim 2020).

Measuring Voter Updating and Behavior

Finally, I measure voter updating and voting behavior in response to provision of information about politician allocation behavior. The study of information revelation and voter updating has spawned a large body of recent field and survey experiments (for a list of these studies, see Bhandari, Larreguy, and Marshall 2021; Dunning et al. 2019; Incerti 2020). I extend a survey experiment fielded in Brazil as reported in Weitz-Shapiro and Winters (2016) and Winters and Weitz-Shapiro (2016) to measure how voter responses to information about politician allocation behavior varies in bureaucratic quality. Some recent work questions the correspondence between voter (respondent) responses to information in survey experiments as opposed to field settings (Boas, Hidalgo, and Melo 2019; Incerti 2020). I contend that the two measure beliefs and actions, respectively. Per the model, the respective effects of information on voter beliefs and vote choice should be different.

The survey experiment thus tests whether voters update beliefs in a manner consistent with the theory. This design affords a clean test of the voter updating mechanism in isolation, separated from the effect of a politician's (strategic) allocation behavior or baseline voter information. It is important to emphasize that the manipulation to voter information refers to politician allocation behavior, not public goods provision (Supplementary Table D6). Consequently, the survey experiment tests the extension of the model described in Figure 2 and Proposition A1 in the Supplementary Material. This test represents a partial equilibrium test of the model extension.

To study equilibrium voter behavior, I consider the phenomenon of incumbency disadvantage to test whether voters anticipate the second-term behavior

---

[13] Among the sample of municipalities audited in the first rounds of randomized audits, audits covered 60% of local budgets.

TABLE 1. Empirical Implications of the Model Relative to the Two Special Cases

EMPIRICAL IMPLICATION	THIS MODEL	EXISTING ALTERNATIVE EXPLANATIONS		EMPIRICAL STRATEGY
		No bureaucratic co-production	**No voter information**	
	$p > 0, 0 < \underline{\theta} < \overline{\theta} < 1$	Assumption: $\underline{\theta} = 0, \overline{\theta} = 1$	Assumption: $p = 0$	
1	*Politician allocations to rents (1−a) weakly decrease in bureaucratic quality (q).*	Allocations to rents are independent of bureaucratic quality (q).	–	Original analysis using measure of rents from Avis, Ferraz, and Finan (2018).
2	*Politicians allocate more or less to rents in their second term (t = 2) than their first term (t = 1). This difference is attenuated to zero at low and high levels of bureaucratic quality.*	Politicians allocate less to rents in their second term than in their first term. This difference is independent of bureaucratic quality.	There is no difference in allocation to rents between first and second terms.	Extension of Ferraz and Finan (2011).
3	*At high levels of bureaucratic quality, a voter's posterior belief (μ) is equivalent to her prior (π) upon receiving a signal that a politician allocated no funds to rents (a = 1).*	Voters update positively (μ > π) in response to a signal that a politician allocated no funds to rents at any level of bureaucratic quality.	Voters do not update (μ = π) in response to a signal of a clean politician at any levels of bureaucratic quality.	Extension of survey experiment reported in Weitz-Shapiro and Winters (2016) and Winters and Weitz-Shapiro (2016).
4	*Incumbency disadvantage does not emerge at low or high levels of bureaucratic quality (q).*	Incumbency disadvantage does not emerge.	Incumbency disadvantage does not emerge.	Extension of Klasnja and Titunik (2017).

Note: "–" indicates that the restricted model does not make a different prediction from the unrestricted model in this article. *For the "no voter information" model, the interpretation in the text is that voters do not observe the signal; the prediction in the cell represents a slightly different interpretation of the same result. Formal motivation of these predictions is included in the Supplementary Material.

of mayors and condition their votes accordingly. Klasnja and Titunik (2017) find evidence of an incumbency disadvantage for Brazilian mayors' parties. Using a close elections regression discontinuity (RD) design, they find that barely elected mayors' parties are less likely to win the subsequent election than barely defeated parties. The present model suggests that we should only observe this incumbency disadvantage in the region of bureaucratic quality in which the incumbent shirks in the second term but not the first. I extend the analysis of Klasnja and Titunik (2017) to examine heterogeneity in the local average treatment effect (LATE) of incumbency on (re)-election as a function of bureaucratic quality. The purpose of this extension is to evaluate whether voters cast their votes in anticipation of second-term shirking where shirking is predicted to occur. Recent critiques by Eggers (2017) and Ashworth, Berry, and Bueno de Mesquita (2021) raise concerns about whether close-elections RD designs can isolate the source of incumbency advantage/disadvantage. In Appendix D5.2 of the Supplementary Material, I show the commensurability of the LATE with voter anticipation of shirking. The difference between these critiques and the present

model is that in the present model, the competence or valence of a challenger is not observed by voters before the election.

Mapping Theory and Meta-Study Design

Empirically, I focus on how observed variation in bureaucratic quality conditions the beliefs and actions of two actors in the model: the politician and voter. Table 1 lays out four reduced-form empirical implications of the equilibria characterized in Proposition 1 (and for equilibrium implication #3, Proposition A1 in the Supplementary Material).[14] Each implication consists of a straightforward extension of one of the constituent studies, as detailed in the right column. I seek to show qualitative correspondence between the four implications that have not previously been systematically connected.

---

[14] Given my argument that the baseline model with a public goods signal has a wider scope than the main extension with a politician allocation signal, I derive equilibrium implications #1, #2, and #4 from the baseline model. However, similar qualitative predictions can be generated from the extension in these cases.

These implications evaluate key claims about variation in the empirical manifestations of electoral accountability across Brazil. Empirical implications #1 and #2 measure the conditions under which politicians allocate budgets to private rents (as opposed to public goods). Empirical implication #3 evaluates how voters update in response to information about politician allocations, which is a necessary condition for voters to re-elect competent types at higher rates. Finally, empirical implications #2 and #4 examine the conditions under which re-election incentives conditions politician allocation behavior (#2) and whether voters anticipate that elimination of these incentives will lead some politicians to shirk in a second term (#4). The model provides clear predictions for how each of these beliefs and behaviors is conditioned or moderated by bureaucratic quality.

Importantly, these empirical implications are distinct from the implications of existing arguments about the function of electoral accountability in Brazil and elsewhere. The two principal departures of my model from these alternatives are that I: (1) incorporate bureaucrats as co-producers of public goods; and (2) assume that voters may have information about the performance of the incumbent politician. It is straightforward to represent the existing accounts as corner cases of the present model in order to tease out the empirical implications of these alternative arguments. One can approximate a model without bureaucratic co-production by assuming that $\underline{\theta} = 0$ and $\overline{\theta} = 1$.[15] This parametric assumption eliminates all pooling and partially pooling equilibria, meaning that the competent type always invests in public goods, while the incompetent type never does so. It is similarly straightforward to evaluate the implications of a (possibly distinct) alternative explanation with completely uninformed voters by assuming that $p = 0$, meaning that the voter does not observe a signal and thus cannot update. This breaks the link between a politician's first-period allocation strategy and their subsequent electoral fortunes. The two sets of parametric assumptions generate the empirical implications reported in columns 2 and 3 of Table 1, which are (generally) distinct from those generated by the model that I advance. These distinct implications allow for evaluation of the current model relative to existing alternatives.

The tests I propose leverage the observable implications of the equilibria that emerge at different levels of bureaucratic quality. This presents a challenge because we do not know the true mapping between q and the empirical measure of bureaucratic quality. For example, it is possible that all municipalities have sufficiently similar bureaucratic quality to ensure that all municipalities fall into one of the equilibria identified in Proposition 1. The inferences from the proposed tests rely upon variation in the underlying equilibria. To this end, I adopt three complementary approaches to

interpretation of the data. First, description of the data and reading of the existing literature provides some guidance on what is plausible. Second, I leverage insights *across* the tests described in Table 1. Because these data come from different sources, qualitative consistency across tests is, in principle, more challenging to achieve but also more informative. Finally and most practically, I am careful to model bureaucratic quality flexibly in regression specifications given the possibility of non-linearities suggested by the model.

Identifying the Effects of Bureaucratic Quality

The empirical implications in Table 1 consist of all-else-equal (comparative static) predictions about how bureaucratic quality conditions or moderates different aspects of the accountability relationship between voters and politicians. However, I do not manipulate levels of bureaucratic quality, nor do I identify a research design which generates plausibly exogenous variation in this variable. This reliance of observational variation in bureaucratic quality raises a number of important theoretical and empirical questions.

The theory implies that the distribution of politician types should also affect various empirical manifestations of electoral accountability relationships between voters and politicians. This distribution is characterized by the share of competent politicians (π), the monitoring rate of competent politicians ($\overline{\theta}$), and the monitoring rate of incompetent politicians ($\underline{\theta}$). This raises two concerns for the evaluation of comparative static predictions. First, if bureaucratic quality covaries with politician competence, then it is possible that observed effects are instead driven by variation in the distribution of politicians and could be mis-attributed to variation in bureaucratic quality. Second, it is possible that simultaneous variation in bureaucratic quality and politician competence strengthens or weakens the all-else-equal effect of bureaucratic quality.

Fortunately, it is straightforward to analyze both concerns theoretically. In Appendix D1.2.1 of the Supplementary Material, I evaluate analogous comparative static implications for each of the parameters of the distribution of politician quality. While some implications (e.g., empirical implication #3 on voter updating) do not distinguish bureaucratic quality from another parameter characterizing the pool of politicians, no other parameter (π, $\overline{\theta}$, or $\underline{\theta}$) produces the set of four empirical implications of variation in bureaucratic quality (q) that is reported in Table 1. I then consider the effect of a compound shock to bureaucratic quality and each of the politician competence parameters in Appendix D1.2.2 of the Supplementary Material. While simultaneous increases in $\overline{\theta}$ and bureaucratic quality strengthen observed effects relative to increases in bureaucratic quality alone,[16] changes in $\overline{\theta}$ alone do not yield the same implications discussed in Table 1

---

[15] Note that this model falls outside parametric restrictions imposed on $\underline{\theta}$ and $\overline{\theta}$ in the model, so this is not precisely a case of the general model.

[16] These complementarities are not general across the support of bureaucratic quality, nor do they emerge on the same intervals across empirical implications.

(column 1). To further alleviate these concerns empirically, in Appendix A2.4 of the Supplementary Material, I analyze the covariance between bureaucratic quality and one measure of politician competence using candidate data from mayoral elections of 2004, 2008, and 2012 (those used in the following analyses). The correlation of these measures is weak ($\rho < 0.11$) and after residualization with the covariate conditioning set, it all but disappears ($\rho < 0.02$).

Empirically, reliance on observational variation in bureaucratic quality raises concerns of confounding of bureaucratic quality and outcomes of interest by one or more unmeasured variables. I employ three strategies to lessen these concerns. First, as described above, all analyses are reported with and without a flexible set of covariates described above. This selection-on-observables design aims to reduce some obvious potential confounds related to local economies, labor markets, and the information environment. However, it is well known that unobserved or unmeasured confounders (e.g., clientelism) could bias our estimates—and thereby inferences—about the moderating effects of bureaucratic quality. To this end, as a second strategy, I employ the sensitivity analysis proposed by Cinelli and Hazlett (2020) to answer an important question: how strong would unmodeled confounding need to be relative to the existing set of covariates to change our substantive findings? Finally, empirical implications #2–#4 leverage additional sources of variation to measure the moderating effect of bureaucratic quality. Design-based strategies—the survey experiment and the RD exercises—greatly reduce the threat that bureaucratic quality and other sources of variation (e.g., voter information) are both confounded by the same unobservables and clarify the contrasts being made.

Estimation

Table 1 presents four empirical implications of the theory. In the first test, I examine the association between politician allocations to rents (corruption) and municipal bureaucratic quality. To examine this relationship, I estimate an OLS regression of the form

$$Y_{msl} = \beta_1 Q_m + \gamma_s + \lambda_l + \delta X_m + \epsilon_{msl}, \qquad (9)$$

where Y_{msl} is the proportion of audited funds allocated to rents in municipality m in state s, as measured in lottery round l. Q_m is the measure of municipal bureaucratic quality and β_1 is the coefficient of interest. γ_s is a vector of state fixed effects and λ_l is a vector of lottery round fixed effects. X_m is a matrix of decile indicators for each of four municipal-level covariates: population, average education, formality rate, and GDP per capita, as well as an indicator for community radio presence. Empirical implication #1 holds that $\beta_1 < 0$: corruption declines in bureaucratic quality. I also estimate this specification with tercile and quartile bins of Q_m given the potential for non-linearity suggested by the model.

Second, I examine whether term differences in the allocation of rents varies in bureaucratic quality. I test this prediction by estimating an OLS regression of the form

$$Y_{msl} = \beta_1 Q_m + \beta_2 \text{Second term}_m$$
$$+ \beta_3 Q_m \text{Second term}_m + \gamma_s + \lambda_l + \delta X_m + \epsilon_{msl}, \qquad (10)$$

where Y_{msl} is again the share of audited funds allocated to rents. Q_m is the measure of municipal bureaucratic quality and the indicator Second term$_m$ captures whether the politician is in her second term in the 2001–04 term (when the audits occurred). Bureaucratic quality, Q_m, is modeled linearly and in quantile bins in different specifications. γ_s is a vector of state fixed effects and λ_l is a vector of lottery round fixed effects. X_m includes the same set of flexible controls for local labor markets. The quantity of interest is $\beta_2 + \beta_3 Q_m$, the marginal effect of being a second term mayor, at a given level of bureaucratic quality. Empirical implication #2 predicts that $\beta_2 + \beta_3 Q_m = 0$ at high levels of Q_m. I also use an RD-based design to decompose selection from shirking effects, enumerated at greater length in Appendix A4 of the Supplementary Material.

Third, using survey experimental data reported by Winters and Weitz-Shapiro (2016) and Weitz-Shapiro and Winters (2016), I estimate how voter updating varies in bureaucratic quality. This experiment was fielded in the context of a nationally representative survey in 140 randomly sampled municipalities. The survey experiment provided respondents with a common vignette about a first-term mayor of a different city seeking re-election (thus a first-term mayor). A control condition provides no information about municipal audit outcomes. Additional text conveying a clean treatment condition indicates that the mayor was found not to have awarded bribes for city contracts.[17] I estimate conditional average treatment effects (CATEs) of the clean signal on voter feelings toward the hypothetical, using OLS regressions of the form

$$Y_{ims} = \beta_1 Q_m + \beta_2 \text{Clean signal}_i + \beta_3 \text{Clean signal}_i Q_m$$
$$+ \gamma_s + \theta X_m + \epsilon_{ims}, \qquad (11)$$

where Y_{ims} is respondent i's assessment of the hypothetical mayor on a seven-point feeling thermometer. The clean signal indicator measures assignment to the

---

[17] Weitz-Shapiro and Winters (2016) and Winters and Weitz-Shapiro (2016) emphasize a different treatment condition in which the mayor was found to be corrupt in the audit. In Supplementary Table A10, I estimate the CATEs of both the clean and corrupt treatment conditions. The clean treatment tests equilibrium implications of the model in any equilibrium where any politician allocates funds to public goods (e.g., $q \geq \tilde{q}_1$), which is consistent with the observed data from empirical implications #1 and #2. Where bureaucratic quality is high, however, the corrupt treatment tests a mix of empirical implications and theoretical assumptions about off-path voter beliefs implied by the intuitive criterion. For this reason, I focus on the clean treatment condition in the main text.

clean vignette. I cluster standard errors at the level at which Q_m is measured: the municipality (indexed by m). The estimator of the relevant CATE is $\beta_2 + \beta_3 Q_m$, which measures updating on the clean signal at a given level of bureaucratic quality.

Finally, I examine rates of incumbency disadvantage to see whether voters vote in a manner consistent with theoretical predictions. Following Klasnja and Titunik (2017), I adopt a close elections RD design. I extend their analysis by estimating the LATE of incumbency at different levels of bureaucratic quality. Specifically, I estimate the conditional LATE given by:

$$\tau_q = \lim_{x\downarrow 0} E[Y_m | Q_m = q] - \lim_{x\uparrow 0} E[Y_m | Q_m = q], \tag{12}$$

where x is the margin of victory (and 0 represents a tied election). The conditional LATEs are measure incumbency disadvantage at different measured levels of bureaucratic quality, q. I use the robust, bias-corrected Calonico, Cattaneo, and Titiunik (2014) RD estimator, fit separately on each (quantile) bin of bureaucratic quality. I maintain the optimal bandwidth from the pooled sample for each subset. The unconditional RD estimand presented in Klasnja and Titunik (2017) can be decomposed into a candidate or party's decision on whether to contest election $t+1$ and citizens' votes given the choices on the ballot (Slough 2023). To this end, I present unconditional LATE estimates in addition to the LATE on party election contestation and a post-treatment estimand that examines vote choice conditional on the incumbent running. While the latter is *not* a standard causal estimand, it enables us to better isolate voter (as opposed to candidate) behavior. I formalize the decomposition of the unconditional LATE and motivate these estimands in Appendix A6.1 of the Supplementary Material.

RESULTS

The four findings follow the empirical implications in Table 1 and assess the plausibility of the theoretical model that I advance. While these tests use distinct samples, treatments, and outcomes, they collectively show that the empirical manifestations of accountability vary systematically in bureaucratic quality in the Brazilian context. I weight the evidentiary value of these tests equally and assess qualitative congruence between the four implications of interest.

Politician Allocations to Rents Decrease in Bureaucratic Quality

In a first test of the theory, I examine the relationship between bureaucratic quality and politician allocations to rents. The theory predicts that politician diversions to rents (as opposed to public goods) are decreasing in bureaucratic quality. Column 1 of Table 2 presents the bivariate regression of the share

of funds allocated to rents $(1-a)$ to rents on bureaucratic quality (q). The estimated coefficient implies that one-standard-deviation increase in bureaucratic quality reduces rents by 1.4 percentage points or 22.5% of the sample mean. The conditional association of bureaucratic quality and mayors' allocations to rents remains substantively similar when including covariates. Given the right-skewed distribution of the share of corrupt spending, I also look at the logged outcome, with substantively similar findings in columns 4–6. Consistent with the model, more flexible specifications of bureaucratic quality in Panels B and C do not indicate non-monotonicity in this relationship. Corrupt spending is concentrated in the lowest quantiles of bureaucratic quality. On average, 7.3% of federal funds in the lowest tercile of municipalities and 7.5% of such funds in the lowest quartile of municipalities are spent in a corrupt manner. These rates drop in higher quantiles of bureaucratic quality.

The finding that rent extraction declines in bureaucratic quality is consistent with the model I advance, but inconsistent with the predictions of a model of accountability that omits bureaucrats (see Table 1). It could, however, be consistent with a world with uninformed voters. A skeptical reader may be concerned that unobserved confounding could be driving this finding. To this end, the sensitivity analysis and benchmarking in Supplementary Tables A7 and A8 reveal that, to explain away the point estimates, an omitted confounder would have to have four to nine times the predictive power of the most predictive covariate and 42%–53% of the predictive power of the full set of covariates and fixed effects. In sum, this analysis provides evidence consistent with the idea that politician allocate more funds to rents (or corruption) when bureaucratic quality is low.

More descriptively, the baseline *levels* of allocation to rents—a mean of 6.2% of audited funds—reported in Table 2 suggest circumscribed corruption, which is consistent with equilibria in which some or all politicians allocate funds to public goods (the converse of rents). Further, Figure 3 plots the density of the corruption (rents) outcome, which suggests that the modal politician did not divert any of the audited federal funds. Indeed, the median politician only allocates 1.9% of these funds to rents. Per the model, circumscribed corruption occurs when bureaucratic quality is high enough to induce at least one type of politician to invest in public goods (the separating equilibrium, the partially pooling equilibrium, or the pooling equilibria where both types invest). This provides suggestive evidence about the distribution of bureaucratic quality in Brazilian municipalities.

Second-Term Shirking Disappears as Bureaucratic Quality Increases

Recall that in the model, the partially pooling and pooling equilibria in which politicians allocate more to public goods in the first period than in the second emerge under two conditions. First, intermediate

TABLE 2. Association between Bureaucratic Quality, *q*, and Allocations to Public Goods, *a*

|	Share of corrupt spending			Log(Share of corrupt spending + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
A. LINEAR BUREAUCRATIC QUALITY MEASURE (*Z*-SCORE)						
Bureaucratic quality	−0.014**	−0.014**	−0.017**	−0.012**	−0.012**	−0.014**
	(0.006)	(0.006)	(0.007)	(0.005)	(0.005)	(0.006)
B. BUREAUCRATIC QUALITY MEASURE TERCILES (RELATIVE TO FIRST TERCILE)						
Bureaucratic quality, tercile 2	−0.009	−0.009	−0.009	−0.007	−0.007	−0.007
	(0.012)	(0.012)	(0.012)	(0.010)	(0.010)	(0.010)
Bureaucratic quality, tercile 3	−0.027**	−0.026*	−0.036**	−0.023**	−0.022*	−0.029**
	(0.012)	(0.014)	(0.018)	(0.010)	(0.011)	(0.014)
C. BUREAUCRATIC QUALITY MEASURE QUARTILE (RELATIVE TO FIRST QUARTILE)						
Bureaucratic quality, quartile 2	−0.009	−0.003	−0.002	−0.006	−0.001	0.000
	(0.015)	(0.015)	(0.015)	(0.012)	(0.013)	(0.012)
Bureaucratic quality, quartile 3	−0.019	−0.021	−0.029*	−0.015	−0.018	−0.024*
	(0.015)	(0.014)	(0.015)	(0.012)	(0.012)	(0.013)
Bureaucratic quality, quartile 4	−0.029**	−0.030*	−0.042**	−0.025**	−0.025*	−0.034**
	(0.014)	(0.016)	(0.021)	(0.012)	(0.013)	(0.017)
State FE		✓	✓		✓	✓
Lottery FE		✓	✓		✓	✓
Demographic controls (decile bins)			✓			✓
Community radio indicator			✓			✓
Outcome range	[0,0.794]	[0,0.794]	[0,0.794]	[0,0.584]	[0,0.584]	[0,0.584]
Outcome mean	0.062	0.062	0.062	0.056	0.056	0.056
Outcome std. dev.	0.10	0.10	0.10	0.085	0.085	0.085
Num. obs.	448	448	448	448	448	448

Note: Funds diverted from public goods are measured as the share of corrupt spending. Heteroskedasticity-robust standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. (See Supplementary Table D2 for suppressed coefficient estimates.)

FIGURE 3. Distribution of the Measure of Rents: Corrupt Spending as a Share of Total Audited Federal Funds
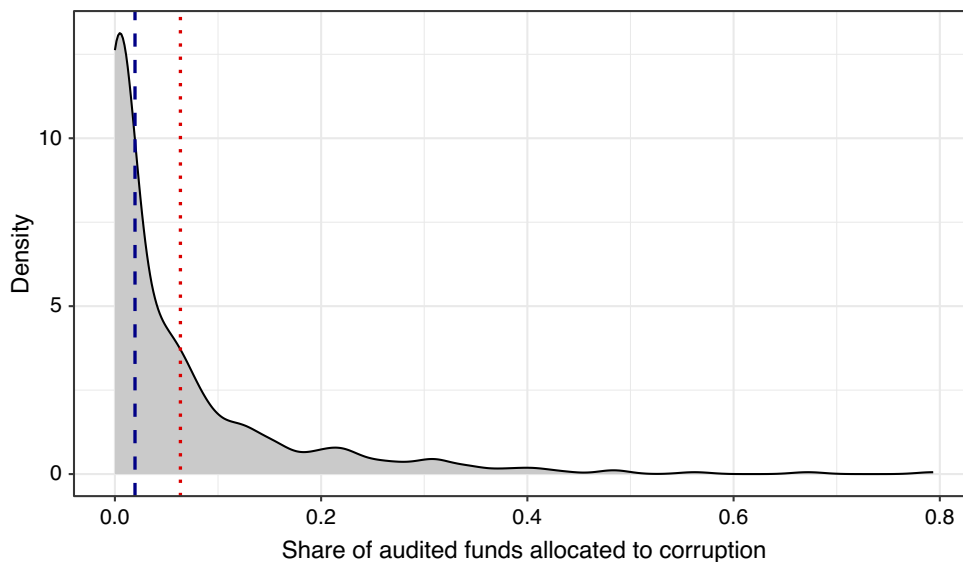
Note: The navy dashed line gives the median, and the red dotted line gives the mean.

bureaucratic quality makes public goods production by the incompetent type of politician inefficient, while public goods production is efficient for the competent type. This inefficiency leads to second-term shirking when electoral incentives are removed. Second, voters are likely enough to observe the signal that an incompetent politician's costly diversion of funds to public goods in the first term can improve their re-election

FIGURE 4. The Marginal Effect of a Second-Term Politician versus a First-Term Politician on Proportion of Audited Funds Allocated to Rents
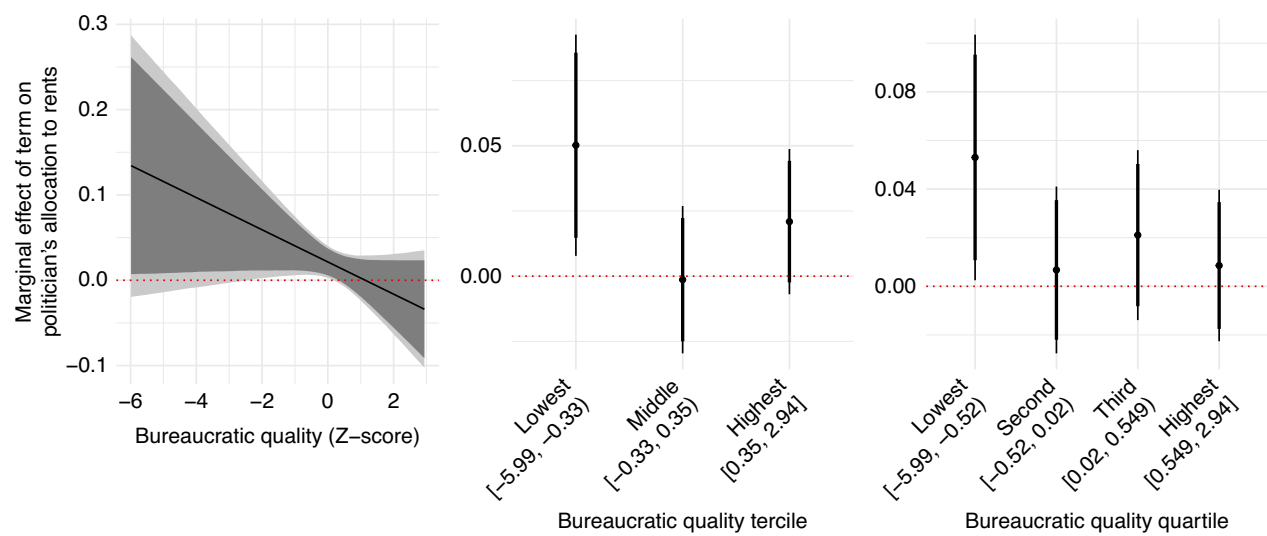
Note: 90% and 95% confidence intervals constructed on heteroskedasticity-robust standard errors.

prospects. Ferraz and Finan (2011) establish that corruption is, on average, higher among second-term mayors than first-term mayors. This finding directly contradicts the predictions of the model without bureaucratic co-production, in which differences in allocations by term are driven only by positive selection in the re-election of mayors. I further examine how this difference between (term-limited) second-term and (non-term limited) first-term mayors varies in bureaucratic quality. Figure 4 shows that second-term shirking —or diversion of funds from public goods—manifests strongly in municipalities with low bureaucratic quality. However, this difference by term disappears as bureaucratic quality increases.

This finding is consistent with the suggestion that the lowest observed levels of bureaucratic quality correspond to the separating equilibrium and higher levels of bureaucratic quality correspond to the pooling equilibrium with public goods allocations from both types in both periods. Most importantly, under the model advanced here, the finding of any difference between first- and second-term mayors suggests that politicians anticipate that voters could learn about their actions (or outcomes). This provides evidence against the case of "no voter information" ($p = 0$). The result in Ferraz and Finan (2011) is sufficient for this conclusion. The finding that this difference varies predictably in bureaucratic quality provides evidence to bolster the plausibility of the present model of accountability.

Serving as a first- versus second-term mayor is clearly not randomly assigned. One may be tempted to rely on an RD design to identify the conditional (local) ATE of term. Under the present model, however, an RD cannot identify the effect of term alone since second-term

mayors are also more likely to be competent than first-term mayors in the relevant parameter space. Yet, under the valence assumptions in the voting model, an RD-like estimator that compares conditional means on both sides of the re-election threshold at different bandwidths provides a way to decompose term effects from politician type. By reducing the sample to close races (setting a small bandwidth), the share of incompetent types in the pool of second-term incumbents should be higher and, thus, we would expect more shirking on average (a larger difference between the first and second terms). By varying the bandwidth in Supplementary Figure A5, I show that this prediction is indeed borne out in the analysis. Consistent with the model's predictions, in samples where incompetent types theoretically represent a larger share of second-term mayors (close elections), second period shirking is more pronounced.

This test and these results provide evidence against an alternative (unmodeled) account in which bureaucrats exert substantial power over politicians as in Raffler (2022). In this alternate account, high-quality bureaucrats prevent politicians from engaging in corruption, whereas weaker, lower-quality bureaucrats do not. If bureaucrats were indeed constraining the politician's allocations in this way, we should not observe second-term shirking. This finding, therefore, provides evidence that is inconsistent with accounts in which bureaucrats are unilaterally driving both allocation decisions and policy outcomes.

These findings on term effects provide important evidence about politician allocation strategies and the role of electoral incentives. The finding of negative second-term effect on politician allocations at low levels of bureaucratic quality (within sample) suggests

15

FIGURE 5.   Effects of Clean Signal in Vignette at Different Levels of Bureaucratic Quality
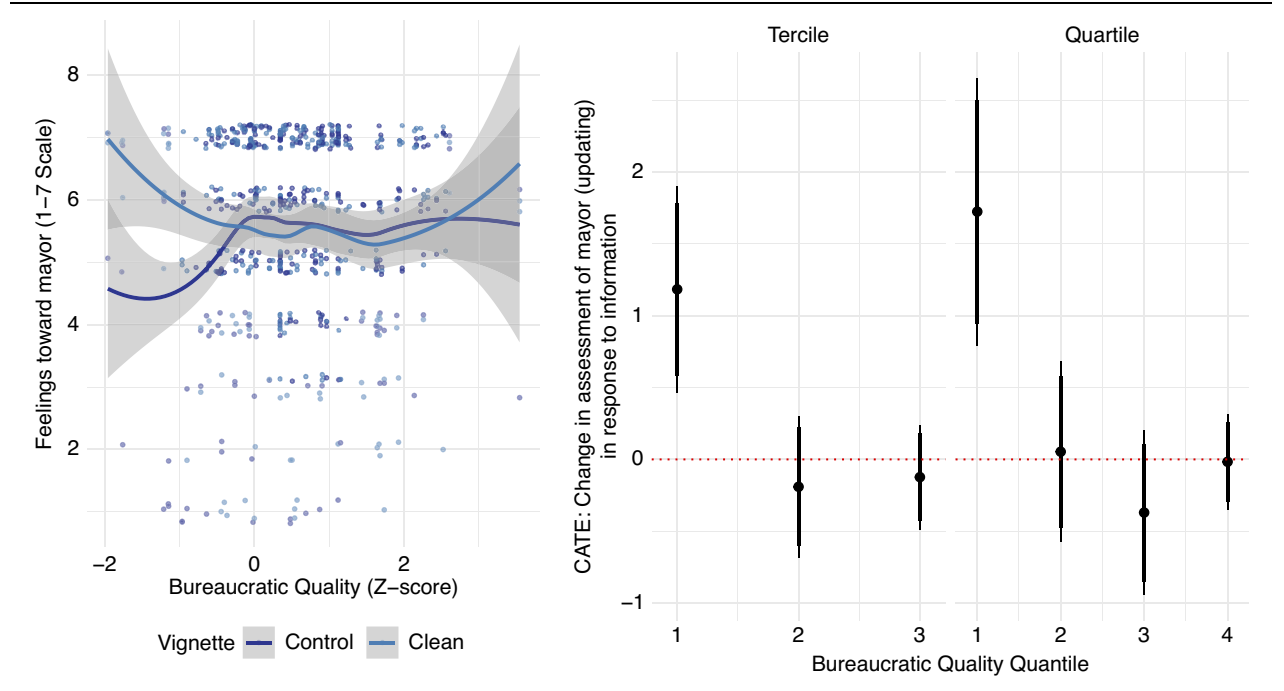
Note: Raw data fitted by Loess (left) and CATEs of the clean signal at different levels of bureaucratic quality (right). All reported CATEs are estimated from OLS regression models with the labor market covariates. Thick and thin lines represent 90% and 95% confidence intervals. Standard errors are clustered at the municipality, the level of aggregation at which bureaucratic quality moderator is measured.

that some politicians internalize electoral incentives that favor investment in public goods in order to win re-election before shirking in a second term. This first-period pooling indicates that (some) politicians anticipate that voters may observe outputs and vote accordingly. However, the *lack* of term effects at high levels of bureaucratic quality—precisely where allocations to rents are lowest—suggests that electoral incentives are not necessary for public goods investments in these contexts. The model posits that efficiency considerations may be sufficient to promote public goods investments with high-quality bureaucracies.

Voters Do Not Update in Response to Revelation of Clean Politicians When Bureaucratic Quality Is High

The third empirical implication of the model holds that, at high levels of bureaucratic quality where competent and incompetent pool and allocate funds to public goods in a first term, voters should not update from a clean signal. In the survey experiment, therefore, the CATE of the clean signal should be attenuated to zero at high levels of bureaucratic quality. The estimates in Figure 5 suggest that respondents only update on a clean signal at low levels of bureaucratic quality (within the sample). In the top two terciles and top three quartiles of bureaucratic quality, respondents' prior (control) and posterior beliefs upon

revelation of a clean mayor are not distinguishable (on average).[18]

Inspection of the data casts doubt on some alternative explanations. First, while the prior belief increases slightly in bureaucratic quality, there is no evidence of a ceiling effect on the CATE, even high levels of bureaucratic quality, as in the left panel of Figure 5. Across quantiles, there is room to observe positive updating on the basis of a clean signal. Second, while demand effects may be of concern with respect to voter updating about corruption, it is harder to explain why demand effects for the clean treatment would diminish with municipal bureaucratic quality. Finally, in this analysis, confounding is a concern with regard to the difference in CATEs at different levels of bureaucratic quality. However, this is not the relevant test of the theory; the CATEs are causally identified by the experimental research design (under standard assumptions).

Thus, the relevant concern is that another feature that drives bureaucratic quality also drives similar patterns of differential updating. To this end, I do not detect variation in these patterns of updating by individual respondent characteristics. In particular, these patterns do not vary detectably in citizen education or

---

[18] The finding of substantial updating on politician in low quantiles is distinct from the null ATE reported in Winters and Weitz-Shapiro (2016).

FIGURE 6. Incumbency Disadvantage of Brazilian Mayors at Different Levels of Bureaucratic Quality
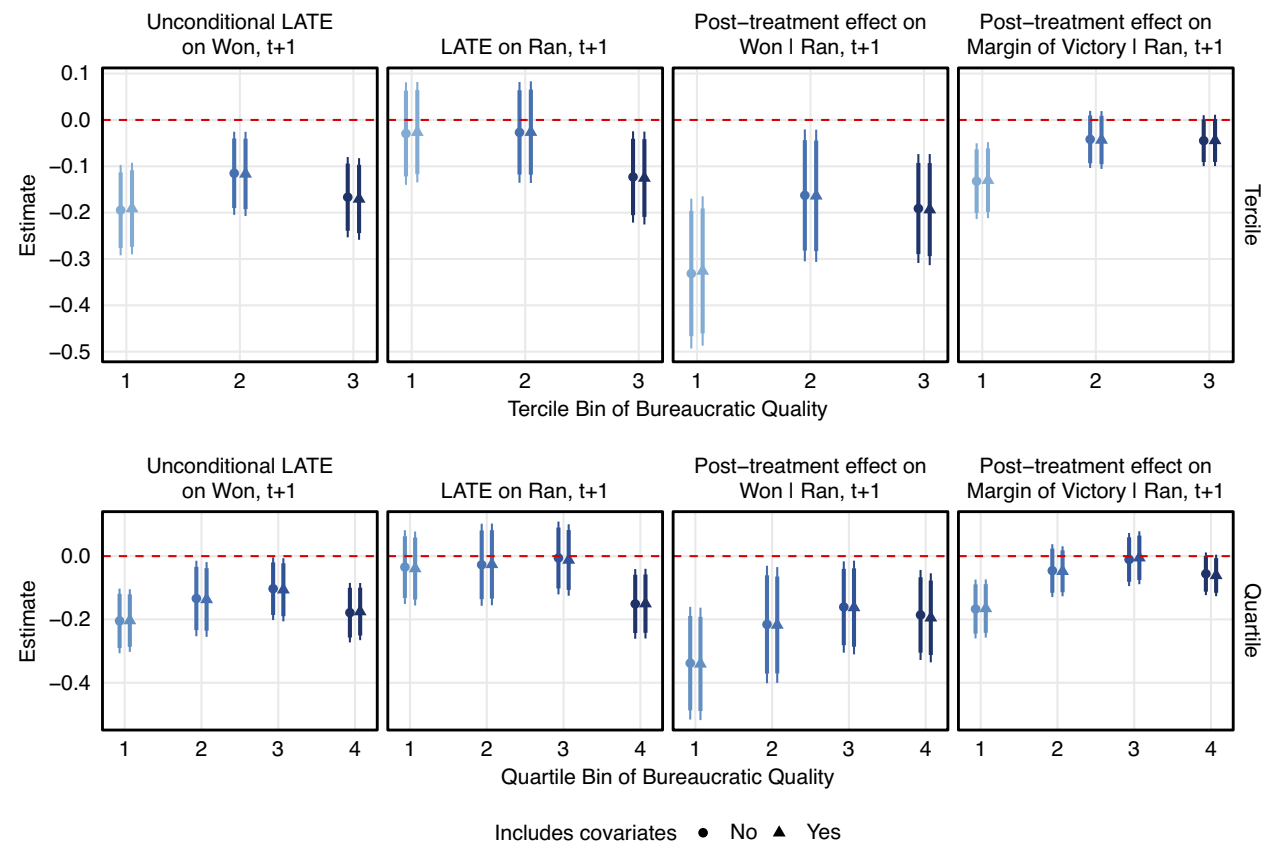
Note: All estimates are estimated using the Calonico, Cattaneo, and Titiunik (2014) bias-corrected regression discontinuity estimator, with 90% (thin) and 95% (thick) confidence intervals constructed on bias-corrected standard errors. The margin of victory outcome in the fourth column of plots is rescaled to a [−1, 1] range.

political knowledge on factual questions, as defined by Weitz-Shapiro and Winters (2016) (Supplementary Figure D6). This helps to allay concerns that differential patterns of updating require sufficient political sophistication or are driven by variation in the educational composition of the electorate that covaries with bureaucratic quality.

In sum, extension of the survey experiment provides evidence that citizens (voters) update differentially as a function of bureaucratic quality. Specifically, good news of a clean record only leads to updating when competent and incompetent types are predicted to take different actions. This suggests that voters internalize expectations about politician behavior and respond to information in light of these expectations. Consistent with the evidence from empirical implication #2 and the descriptive analysis of levels of corruption in the audit data, this occurs at low levels of bureaucratic quality within the sample of Brazilian municipalities. This variation in voter learning is consistent with Bayesian updating and supports the idea that, conditional on receiving information, voters can process this information rationally. Where information allows voters to distinguish competent from incompetent types, this

updating is a precondition for retaining competent politicians at higher rates (positive selection of second-period politicians).

Incumbency Disadvantage and Equilibrium Voter Behavior

Under the model, incumbency disadvantage emerges when voters anticipate that incompetent incumbents allocate funds to public goods in the first period but shirk in the second period. As a result, voters prefer a first-term incumbent of either type to a second-term incompetent type and vote to re-elect incumbent at lower rates. This phenomenon should emerge only in regions of the parameter space where a first-term, incompetent incumbent can induced to pool with the competent type by allocating the budget to public goods in order to improve their re-election chances.

I report the results on incumbency disadvantage at different levels of bureaucratic quality in Figure 6. The left-most panels in each row of plots present estimates of the *unconditional* incumbency disadvantage reported in Klasnja and Titunik (2017). Two observations are of note. Inconsistent with both alternative

explanations in Table 1, we observe incumbency disadvantage in all subgroups, as evidenced by the negative and statistically significant LATE estimates. The findings in the previous tests suggests that we should expect incumbency disadvantage to emerge only at low levels of bureaucratic quality in the present sample. While the point estimate is largest in magnitude in this sample, the differences are only marginally statistically significant.[19] Perhaps more concerning, larger incumbency disadvantage emerges at low and high levels of bureaucratic quality (in sample).

To understand this tension between all model predictions and the RD results, following the analysis in Appendix A6.1 of the Supplementary Material, I decompose this LATE into the LATE on running in time $t + 1$ and a post-treatment estimand on winning (conditional on running) in $t + 1$. These estimates are reported in the second and third columns of Figure 6. This decomposition suggests that incumbent parties run at lower rates in the highest tercile or quartile of bureaucratic quality, a phenomenon that is outside the scope of the model. However, conditional on having a choice to vote for the incumbent party, voters punish incumbent parties to a greater extent where bureaucratic quality is lowest and second-period shirking is most likely as in Figure 4. The fourth column reports a related post-treatment estimand on vote share, supporting the result that voters punish incumbents more strongly at low levels of bureaucratic quality (in sample).

The evidence on incumbency advantage is more ambiguous than the other tests. Specifically, incumbency disadvantage arises across all levels of bureaucratic quality. Nevertheless, the decomposition of this estimand suggests that when the option is available to voters, voters vote against incumbents at higher rates in regions of bureaucratic quality consistent with second-term politician shirking. This increases the magnitude of the incumbency disadvantage. These findings provide qualified evidence that, in equilibrium, voters anticipate second-term shirking in precisely the regions of bureaucratic quality in which it is most likely to emerge. This suggests that (some) voters understand that electoral incentives can motivate first-term politician investment in public goods.

ADDITIONAL APPLICATION

The empirical analysis of mayoral accountability in Brazil relies on a novel theoretically structured meta-study to understand how bureaucratic quality produces heterogeneity in electoral accountability relationships across municipalities. I find qualitatively consistent patterns of heterogeneity in the equilibria that present as a function of bureaucratic quality, even in different samples of municipal elections with different treatments and outcomes of interest.

This form of meta-study can also be used to examine heterogeneity (or conflicting findings) in studies from more distinct settings. In section A7 of the Supplementary Material, I provide a second theoretically structured meta-study that examines when revelation of good and bad news affect voter assessment of incumbents across 11 experiments in 8 countries. While case selection (where experiments are conducted) and measurement of bureaucratic quality across contexts are more challenging in this study than in the Brazilian context, the analysis provides suggestive evidence that voter learning increases in national bureaucratic quality within the sample of low- to middle-income democracies where these experiments have been conducted. This application highlights the range and flexibility of the theoretically structured meta-study as a strategy for accumulating evidence.

CONCLUSION

I argue that variation in bureaucratic quality can help us to systematize widespread heterogeneity and conflicting findings in the study of voter–politician accountability relationships. Where bureaucratic quality is very low, politicians know that their public goods investments are not likely to materialize, so both competent and incompetent politicians instead invest in private rents. In this environment, voters cannot and should not use public goods outcomes to learn about politician quality. At intermediate levels of bureaucratic quality, competent—but not incompetent—politicians can efficiently enact public goods through the bureaucracy and voters can and should use public goods provision to learn about incumbents and retain competent politicians. At somewhat higher levels of bureaucratic quality, this voter learning can induce incompetent types to make costly public goods investments during a first-term to improve their electoral fortunes. Finally, high levels of bureaucratic quality render public goods investments profitable for both competent and incompetent politicians, spurring pooling in investments in public goods from all politicians. Because public goods outputs still depend on a bureaucrat to produce, voters can learn about politician competence and retain competent politicians if they observe public goods in these settings.

This article opens two agendas for future research. First, this article lays the groundwork for forging new connections between state organization and electoral accountability. While recent contributions such as Nathan (2016), Martin and Raffler (2021), and Gottlieb (2024) show how various manifestations of limited state capacity influence voter–politician relationships, I provide the first theory and evidence of how accountability relationships vary systematically across all levels of bureaucratic quality, greatly widening the scope conditions for this line of investigation. Quality is one of many characteristics of bureaucracies—and thereby states—as organizations. Future empirical and theoretical work should explore how other aspects of

---

[19] For the tercile measure, $p < 0.07$ for the tercile 1 and tercile 2 difference; for the quartile measure, $p < 0.18$ for the quartile 1 and quartile 2 difference and $p < 0.05$ for the quartile 1 and quartile 3 difference.

bureaucratic organizations, incentives, actions, or outputs condition variation in electoral politics across contexts. Such work will enrich our understanding of the relationship between state capacity and democratic politics in a wide range of contexts.

Second, the research design that I introduce, a theoretically structured meta-study, opens new avenues for cumulative research in the social sciences. The design shows how theory and empirics can be jointly harnessed to promote evidence accumulation in literatures with substantial bodies of empirical evidence. However, like most multi-method studies, the current results rely on qualitative congruence between findings from discrete studies. Future development of this design should propose more formal means of integration of the statistical tests across constituent studies/estimates. This work will necessitate further substantive consideration of how to weigh different empirical implications and pieces of evidence related to a common theory. Such work will deepen social scientists' ability to accumulate evidence on accountability and beyond.

SUPPLEMENTARY MATERIAL

DATA AVAILABILITY STATEMENT

ACKNOWLEDGMENTS

FUNDING STATEMENT

CONFLICT OF INTEREST

The author declares no ethical issues or conflicts of interest in this research.

ETHICAL STANDARDS

The author affirms this research did not involve human subjects.

REFERENCES

Acemoglu, Daron, Camilo García-Jimeno, and James A. Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105 (8): 2364–409.

Achen, Christopher H., and Larry M. Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton, NJ: Princeton University Press.

Adida, Claire L., Jessica Gottlieb, Eric Kramon, and Gwyneth McClendon. 2020. "When Does Information Influence Voters? The Joint Importance of Salience and Coordination." *Comparative Political Studies* 53 (6): 851–91.

Arias, Eric, Horacio Larreguy, John Marshall, and Pablo Querubin. 2022. "Priors Rule: When Do Malfeasance Revelations Help or Hurt Incumbent Parties?" *Journal of the European Economic Association* 20 (4): 1433–77.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S. Lee. 2020. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *American Economic Review* 110 (5): 1355–94.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15: 183–201.

Ashworth, Scott, Christopher R. Berry, and Ethan Bueno de Mesquita. 2021. *Theory and Credibility: Integrating Theoretical and Empirical Social Science.* Princeton, NJ: Princeton University Press.

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Fridenberg. 2018. "Learning about Voter Rationality." *American Journal of Political Science* 62 (1): 37–54.

Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126 (5): 1912–64.

Berwick, Elissa, and Fotini Christia. 2018. "State Capacity Redux: Integrating Classical and Experimental Contributions to an Enduring Debate." *Annual Review of Political Science* 21: 71–91.

Besley, Tomithy, Robin Burgess, Adnan Khan, and Guo Xu. 2022. "Bureaucracy and Development." *Annual Review of Economics* 14: 397–424.

Bhandari, Abhit, Horacio Larreguy, and John Marshall. 2021. "Able and Mostly Willing: An Empirical Anatomy of Information's Effect on Voter-Driven Accountability in Senegal." *American Journal of Political Science* 67 (4): 1040–66.

Boas, Taylor C., and F. Daniel Hidalgo. 2011. "Controlling the Airwaves: Incumbency Advantage and Community Radio in Brazil." *American Journal of Political Science* 55 (4): 869–85.

Boas, Taylor C., F. Daniel Hidalgo, and Marcus André Melo. 2019. "Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil." *American Journal of Political Science* 63 (2): 385–400.

Brierly, Sarah. 2020. "Unprincipled Principals: Co-Opted Bureaucrats and Corruption in Local Government in Ghana." *American Journal of Political Science* 64 (2): 209–22.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs." *Econometrica* 82 (6): 2285–326.

Calvo, Ernesto, and Maria Victoria Murillo. 2004. "Who Delivers? Partisan Clients in the Argentine Electoral Market." *American Journal of Political Science* 48 (4): 742–57.

Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 1902 (2): 179–221.

Chong, Alberto, Ana de la O, Dean Karlan, and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification." *Journal of Politics* 77 (1): 51–77.

Cinelli, Carlos, and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society B* 82 (1): 39–67.

Cruz, Cesi, Philip Keefer, and Julien Labonne. 2018. "Buying Informed Voters: New Effects of Information on Voters and Candidates." *The Economic Journal* 131 (635): 1105–34.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128 (3): 1169–218.

Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Craig McIntosh, and Gareth Nellis, eds. 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I*. New York: Cambridge University Press.

Eggers, Andrew C. 2017. "Quality-Based Explanations of Incumbency Effects." *Journal of Politics* 79 (4): 1315–28.

Enríquez, José Ramón, Horacio Larreguy, John Marshall, and Alberto Simpser. 2023. "Mass Political Information on Social Media: Facebook Ads, Electorate Saturation and Electoral Accountability in Mexico." Working Paper.

Ferrali, Romain, and Galileu Kim. 2020. "Are Anti-Corruption Audits Effective? Evidence From Brazilian Municipal Bureaucracies." Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract˙id=3499117.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123 (2): 703–45.

Ferraz, Claudio, and Frederico Finan. 2011. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." *American Economic Review* 101 (4): 1274–311.

Foarta, Dana. 2023. "How Organizational Capacity Can Improve Electoral Accountability." *American Journal of Political Science* 67 (3): 776–89.

Fowler, Anthony, and Andrew B. Hall. 2018. "Do Shark Attacks Influence Presidential Elections? Reassessing a Prominent Finding on Voter Competence." *Journal of Politics* 80 (4): 1423–37.

Fowler, Anthony, and B. Pablo Montagnes. 2015. "College Football, Elections, and False-Positive Results in Observational Research." *Proceedings of the National Academy of Sciences* 112 (45): 13800–4.

Geddes, Barbara. 1994. *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley: University of California Press.

Gonçalves, Sónia. 2013. "The Effects of Participatory Budgeting on Municipal Expenditures and Infant Mortality in Brazil." *World Development* 53: 94–110.

Gottlieb, Jessica. 2024. "How Economic Informality Constrains Demand for Programmatic Policy." *American Journal of Political Science* 68 (1): 271–88.

Grindle, Merilee S. 2012. *Jobs for the Boys: Patronage and the State in Comparative Perspective*. Cambridge, MA: Harvard University Press.

Healy, Andrew J., Neil Malhotra, and Cecilia Hyunjung Mo. 2010. "Irrelevant Events Affect Voters' Evaluations of Government Performance." *Proceedings of the National Academy of Sciences* 107 (29): 12804–9.

Huber, John D., and Michael M. Ting. 2021. "Civil Service and Patronage in Bureaucracies." *Journal of Politics* 83 (3): 902–16.

Humphreys, Macartan, and Jeremy M. Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda—Preliminary Analysis." IGC Working Paper S-5021-UGA-1.

Incerti, Trevor. 2020. "Corruption Information and Vote Share: A Meta-Analysis and Lessons for Experimental Design." *American Political Science Review* 114 (3): 761–74.

Izzo, Federica, Torun Dewan, and Stephane Wolton. 2022. "Cumulative Knowledge in the Social Sciences: The Case of Improving Voters' Information." Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3239047.

Klasnja, Marko, and Rocío Titiunik. 2017. "The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability." *American Political Science Review* 111 (1): 129–48.

Larreguy, Horacio, John Marshall, and James M. Snyder Jr. 2020. "Publicizing Malfeasance: When the Local Media Structure Facilitates Electoral Accountability in Mexico." *Economic Journal* 130 (631): 2291–327.

Li, Christopher M., Greg Sasso, and Ian Turner. 2023. "Hierarchical Control." Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4444412.

Martin, Lucy, and Pia J. Raffler. 2021. "Fault Lines: The Effects of Bureaucratic Power on Electoral Accountability." *American Journal of Political Science* 65 (1): 210–24.

Nathan, Noah L. 2016. "Does Participation Reinforce Patronage? Policy Preferences, Turnout and Class in Urban Ghana." *British Journal of Political Science* 49 (1): 229–55.

Raffler, Pia. 2022. "Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from a Dominant Party Regime." *American Political Science Review* 116 (4): 1443–59.

Rauch, James E. 1995. "Bureaucracy, Infrastructure, and Economic Growth: Evidence from U.S. Cities during the Progressive Era." *American Economic Review* 85 (4): 968–79.

Seabra, Sérgio N. 2018. "How Auditing Can Be Directed against Corruption: The Case of Brazil's Municipality-Facing Auditing Program." Working Paper, University of Oxford. https://www.bsg.ox.ac.uk/sites/default/files/2018-07/BIP%20Working%20Paper-S%20Seabra.pdf.

Slough, Tara. 2023. "Phantom Counterfactuals." *American Journal of Political Science* 67 (1): 137–53.

Slough, Tara. 2024. "Replication Data for: Bureaucratic Quality and Electoral Accountability." Harvard Dataverse. Dataset. https://doi.org/10.7910/DVN/KKEVEA.

Slough, Tara, and Scott A. Tyson. 2023. "External Validity and Meta-Analysis." *American Journal of Political Science* 67 (2): 440–55.

Toral, Guillermo. 2023. "How Patronage Delivers: Political Appointments, Bureaucratic Accountability, and Service Delivery in Brazil." *American Journal of Political Science*. https://doi.org/10.1111/ajps.12758.

Varjão, Carlos. 2019. "The Role of Local Media in Selecting and Disciplining Politicians." Working Paper, Stanford University.

Weitz-Shapiro, Rebecca, and Matthew S. Winters. 2016. "Can Citizens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil." *Journal of Politics* 79 (1): 60–74.

Winters, Matthew S., and Rebecca Weitz-Shapiro. 2016. "Who's in Charge Here? Direct and Indirect Accusations and Voter Punishment of Corruption." *Political Research Quarterly* 69 (2): 207–19.

Yazaki, Yukihiro. 2018. "The Effects of Bureaucracy on Political Accountability and Electoral Selection." *European Journal of Political Economy* 51: 57–68.